Collaborative Study of the Determination of Tar and Nicotine in Cigarette Smoke

By C. L. OGG[1] and E. FRED SCHULTZ, JR.[2] (Agricultural Research Service, U.S. Department of Agriculture)

Ten laboratories participated in a 2-phase collaborative study of a method for tar and nicotine in cigarette smoke. In one phase, 200 monitor cigarettes were smoked, while in the second phase, 60 cigarettes from each of 5 pairs of samples were smoked over a period of 2 weeks according to a random port × sample design. A total of over 9000 measurements was reported for weight, number of puffs, total particulate matter, water, nicotine, and tar. Statistical analysis of the results showed the significant variables to be within- and among-laboratory differences and a laboratory × material interaction. Agreement between laboratories was good for cigarettes delivering less than 20 mg tar and 1 mg nicotine, but poor for samples with higher tar and nicotine deliveries. Further study of the method is recommended to improve its performance with cigarettes having high tar and nicotine deliveries.

A 2-phase collaborative study of the method for tar[3] and nicotine[4] in cigarette smoke was carried out during the year. The method was the same as that used in a previous study (1) with 2 exceptions. First, the moisture content of the total particulate matter was determined this year and, second, the tar content was calculated by subtracting the moisture and nicotine contents from the observed total particulate matter (TPM). Except for the butt length and the solvent used to extract moisture and nicotine from the filter pads holding the TPM, the method tested is essentially the same as that used by the Federal Trade Commission (2), which was based on the earlier report (1) of the Associate Referee. The 30 mm butt length specified is the same as that used in our earlier study and, like all the smoking parameters, it is based on the habits of the "average" smoker, i.e., it is the average butt length. Two solvents, dioxane and isopropanol, are commonly used for extracting moisture and nicotine from the TPM caught by the filter pads. Isopropanol was chosen for this study because it is less toxic than the dioxane used in the Federal Trade Commission method.

The Analytical Methods Committee of the Tobacco Chemists' Conference, working with the Biometrical Services Staff, Agricultural Research Service, prepared an experimental design for the study which included 2 phases. The first phase was the determination of tar and nicotine in a monitor cigarette, using a sample of 200 cigarettes, half to be smoked during one run (20 ports of 5 each) on one day and the other half to be smoked another day. The second phase required that 60 cigarettes from each of 5 pairs of samples be smoked according to a random port × sample design designated by the Biometrics staff. This required that 100 cigarettes (2 ports for each sample) be smoked during each of 3 days one week and repeated on 3 days of a following week. Data from both phases of the study were analyzed by the Biometrics staff.

One of the 5 pairs of samples consisted of a sample from the old supply of monitor cigarettes and a sample from the new supply. The other 4 pairs were samples from lots of cigarettes chosen to represent the range of tar and nicotine of interest. These ranged from a light tar and nicotine, filtered and ventilated cigarette up to a heavy tar and nicotine, non-filtered cigarette. The first sample of such a pair consisted of cigarettes taken from the production line at a time when the process was judged to be in control. The second consisted of cigarettes taken at a later date when the process was again judged to be in control, but it was thought probable that the hoppers of tobacco had been refilled, the roll of paper had been changed, and the supply of filters might be from a different lot. Thus samples were taken under somewhat more uniform conditions than would prevail in sampling the commercial production of a brand of cigarettes. The pairs of samples can be taken as the pairs of *similar* materials suggested by Youden (3) for the unit blocks of the

collaborative method he proposed. There is one distinct difference, however; the random taking of these samples to represent the production line they were taken from relieves one of the responsibility of deciding whether they are *sufficiently similar* to be treated by the statistical method proposed by Youden. Further, this offers the opportunity to determine estimates of underlying random error appropriate for judging whether possible expression of laboratory × material interactions represents more than unfortunate accidents of sampling.

METHOD

See *JAOAC* **52**, 458–462 (1969), with the following modifications:

Reagent

Isopropanol-ethanol solution.—Reagent grade isopropanol-ethanol solution, in the range 1000 + 2 to 1000 + 6, depending on sensitivity of GLC apparatus. (Use throughout method instead of dioxane-isopropanol solution.)

Apparatus

(a) *Smoking machine and Cambridge filter assembly.*—See *JAOAC* **52**, 458–462 (1969). Check puff volume at each port with smoke collection trap in system at least twice a day or as often as each run if leaks are encountered frequently. Measure puff volume with soap bubble flowmeter having 0.1 ml graduation intervals. Avoid excess drainage on walls of flowmeter when making measurement. Filter holder must be connected as closely as possible to puffing device so that volume between two is held to absolute minimum.

(b) *Gas chromatograph.*—Recommended column and conditions: 3′ × ¼″ stainless steel column packed with 120–150 mesh Porapak Q (Waters Associates, 61 Fountain St., Framingham, Mass. 01701); column 190°C; injection port 240°C; thermal conductivity detector 200°C; helium carrier gas (humidified by passing gas over tube containing water as directed by Sloan and Sublett, *Tobacco Sci.* 9, 70–74 (1965)), flow rate ca 100 ml/min. Adjust sensitivity so that 4 mg water/ml solvent (40 μg/10 μl) gives ca full scale response on 1 mv recorder with bridge current of 210 ma. Column dimensions, temperatures, flow rate, and bridge current may be adjusted to give maximum resolution, sensitivity, and repeatability in minimum time.

Sample Treatment

See *JAOAC* **52**, 458–462 (1969), with the following modification: Mark each cigarette 30 mm from butt end with soft lead pencil or other suitable device without puncturing paper. If filter tip plus overlay exceeds 27 mm, mark cigarette 3 mm beyond overlay and note butt length.

Determination

Particulate matter.—Collect smoke from 5 cigarettes on each filter pad and disconnect filter assembly, gently wipe, and weigh to nearest 0.2 mg. Record gain in weight of filter assembly and total number of puffs and save smoke sample for water and nicotine analyses. Repeat this determination until specified number of cigarettes have been smoked and data recorded.

Moisture.—Remove filter pad from holder and place in 25 ml Erlenmeyer flask or 30 ml serum bottle. Wipe filter holder with ¼ of unused filter pad and place in flask or bottle. Add 10 ml isopropanol-ethanol solution, stopper with puncture-type, serum bottle stopper, and shake 20 min. Withdraw 5 μl sample with 10 μl syringe and inject into gas chromatograph. Determine water to ethanol peak height or area ratio. Subtract peak height or area ratio obtained from blank run using 1¼ conditioned filter pad. Determine mg water in sample as follows: Construct calibration curve by plotting mg water added versus peak height or area ratio of water to standard ethanol. Use 3 or 4 points representing 0–20 mg water per 10 ml solvent standard solution. Determine slope of calibration curve and multiply corrected sample ratio by 1/slope to obtain mg water in sample. Slope of calibration curve should be determined daily.

Nicotine (using modified Kjeldahl still).—Transfer solution and filter pad from moisture analysis step to 500 ml Kjeldahl flask. Rinse bottle with 50 ml 0.1*N* HCl and add this to flask. (Alternatively, 4 ml aliquot of solution may be transferred by pipet to Kjeldahl flask containing 50 ml 0.1*N* HCl.) Fit flask for steam distillation with steam inlet tube, spray trap, and condenser. Steam distill acid solution 10–15 min, keeping volume approximately constant by applying more heat. Discard condensate. Stop steam distillation, place 500 ml flask containing 25 ml HCl (1 + 11) (or, if 4 ml aliquot was used, use 250 ml volumetric flask containing 10 ml HCl (1 + 9)) under condenser with condenser tip dipping into acid solution, add 25 ml sodium hydroxide-salt solution to distillation flask, and connect immediately. Keeping volume in distilling flask between 75 and 100 ml, rapidly steam distill until volume of distillate is ca 450 ml (or 225 ml for 4 ml sample); dilute to volume and mix. Determine A of distillate at 236, 259, and 282 nm against blank of 0.05*N* HCl, using 1 cm cells. Calculate total weight of nicotine in smoke sample as follows:

$$A'_{259} = A \text{ of nicotine corrected for background} = 1.059\,[A_{259} - \tfrac{1}{2}\,(A_{236} + A_{282})]$$

Total mg nicotine per port = $A'_{259} \times$ ml distillate/$(a \times b)$, where a = absorptivity of nicotine in 0.05N HCl and b = cell length. If 4 ml aliquot is used for analysis, multiply answer by 2.5.

$a = A/(c \times b)$, where A = absorbance at 259 nm and c = mg/ml of pure nicotine in 0.05N HCl. Purify nicotine by repeated distillation until physical constants reach constant values which agree with those for pure nicotine.

Nicotine (using Griffith still).—Transfer 4 ml aliquot from moisture analysis step to Griffith still containing 1 ml 1.0N HCl. Keeping volume approximately constant, rapidly steam distill acid solution until volume of distillate is ca 100 ml. Discard distillate. Turn off steam, place 250 ml volumetric flask containing 10 ml HCl (1 + 9) under condenser with condenser tip dipping into acid solution, and add 5 ml sodium hydroxide-salt solution to distillation flask. Keeping volume in flask approximately constant, rapidly distill about 225 ml, dilute to volume, and mix. Proceed as for modified Kjeldahl still, beginning "Determine A of distillate ...".

Tar.—Calculate mg tar per port as mg TPM minus mg water minus mg nicotine, all per port.

Results and Discussion

Results are reported for 6 measured characteristics: (1) weight (g) of 5 cigarettes, (2) number of puffs per 5 cigarettes, (3) mg wet weight of total particulate matter (TPM) per 5 cigarettes, (4) mg water per 5 cigarettes, (5) mg nicotine per 5 cigarettes, and (6) mg tar per 5 cigarettes. Results are reported on a "per port" basis, i.e., as total quantities per 5 cigarettes, because that is the way the data were taken. Any per port result except variances (mean, difference, standard deviation, confidence limit, etc., but not squared measures) can be converted to a per cigarette result by dividing by 5. Mean squares and components must be divided by $5^2 = 25$.

Table 1 reports average daily values for the 6 characteristics as obtained for the monitor cigarettes by the participating laboratories in Study 1. Entries in Table 1 result from evaluating the monitor cigarette on all 20 ports of a machine at a single run and averaging the 20 results into a single value. Mean squares from analyses of variance of the results for nicotine and tar in Study 1 are reported by laboratories in Tables 2 and 3, respectively. Mean squares from overall or combined analyses of variance of 8 laboratories in Study 1 are reported for all 6 measured characteristics in Table 4. Laboratory 25 was omitted from the combined analyses because not all variables were reported. Laboratory 13 was omitted because of aberrant results suspected of being due to a known excessive rate of air flow in the smoking room.

Average results in mg nicotine per 5 cigarettes per sample and average results in mg tar per 5 cigarettes per sample for Study 2 are displayed in Figs. 1 and 2, respectively, for the 5 lots.

Figures 1 and 2 are 2-sample charts as described by Youden (3), representing the 5 pairs of similar materials constructed from the pairs of samples drawn from the 5 lots of Study 2. Each point is an average of results from 12 ports (2 ports on each of 3 days in each of 2 weeks) instead of single results. Use of such averages (as opposed to use of single values as described by Youden) tends to reduce the expression of internal precision or within-laboratory variability relative to systematic lab error or between-laboratory variability.

Average mg nicotine per 5 cigarettes and mg tar per 5 cigarettes obtained for the 5 lots of Study 2 are reported by laboratories in Tables 5 and 6, respectively. Each reported value is the average of results from 24 ports (2 ports on each of 3 days in each of 2 weeks for each of 2 samples). Mean squares from analyses of variance of the results for tar and nicotine in Study 2 are reported by laboratories in Tables 7 and 8, respectively. Mean squares from overall or combined analyses of variance of 8 laboratories for all 6 measured characteristics in Study 2 are reported in Table 9. Laboratories 25 and 13 were omitted from these combined analyses also, as previously discussed for Study 1.

Estimates of the components of random variation affecting the results in Studies 1 and 2 were obtained from the mean squares of Tables 4 and 9, respectively, and are reported in Tables 10 and 11, respectively, for all 6 measured characteristics. The components are also expressed as their respective percentage parts of the random variation in a single observation attributable to the evaluation procedure, i.e., the variability of the *samples* presented to the procedure is *not* included in the percentage estimates.

During the conduct of the studies Laboratory 13 noted that the rate of air flow through the

Table 1. Average values of all 6 measured characteristics[a] of the monitor cigarettes by laboratories and days, Study 1

Lab	Day	Cigarette Wt, g	No. of Puffs	Wet Wt TPM, mg	Water, mg	Nicotine, mg	Tar, mg
1	1	5.50	44.3	108.4	14.7	5.50	88.19
	2	5.47	44.0	108.4	14.1	5.42	88.92
	Av.	5.48	44.2	108.4	14.4	5.46	88.56
2	1	5.38	43.9	118.6	17.3	6.47	94.85
	2	5.40	42.6	114.3	17.1	6.39	90.74
	Av.	5.39	43.2	116.4	17.2	6.43	92.80
4	1	5.45	46.2	111.7	15.1	5.84	90.76
	2	5.39	44.6	107.8	13.0	5.85	88.96
	Av.	5.42	45.4	109.8	14.0	5.84	89.86
8	1	5.43	42.8	112.2	17.2	5.11	89.87
	2	5.32	42.2	108.5	17.6	4.98	85.92
	Av.	5.38	42.5	110.4	17.4	5.04	87.90
13	1	not measured	40.8	101.8	10.6	5.65	85.50
	2		42.7	106.6	11.4	5.74	89.33
	Av.		41.8	104.2	11.0	5.70	87.42
15	1	5.35	40.6	109.9	14.7	6.01	89.24
	2	5.34	40.4	110.3	16.6	5.84	87.80
	Av.	5.34	40.5	110.1	15.6	5.92	88.52
19	1	5.39	45.2	111.7	13.6	5.85	91.72
	2	5.34	43.3	106.7	14.0	5.79	86.90
	Av.	5.36	44.2	109.2	13.8	5.82	89.31
21	1	5.42	43.8	110.1	15.7	5.96	88.40
	2	5.42	43.6	109.1	16.5	5.97	87.16
	Av.	5.42	43.7	109.6	16.1	5.96	87.78
24	1	5.45	43.4	112.8	15.4	5.95	91.44
	2	5.42	43.0	112.0	15.1	5.97	91.00
	Av.	5.44	43.2	112.4	15.2	5.96	91.22
25	1	5.43	43.4	105.6	not measured	6.00	not measured
	2	5.41	44.3	107.3		6.20	
	Av.	5.42	43.8	106.4		6.10	
Overall average[b]		5.40	43.4	110.8	15.5	5.80	89.49

[a] All characteristics evaluated as averages per 5 cigarettes.
[b] Omitting Laboratories 13 and 25.

smoking room was more than at other laboratories. Their results indicated fewer puffs per cigarette with recovery of smaller amounts of wet total particulate matter, water, nicotine, and tar—exactly the kind of results to expect if ambient air movement fanned the burning cone and burned more tobacco between puffs than was burned at the other laboratories. Results from Laboratory 13 were therefore eliminated from the combined analyses. Results in Tables 2, 3, and 8 might indicate that Laboratory 21 was more variable internally than the others, but the results in Table 7 would indicate that Laboratory 21 was no more variable than the others. Examination of Fig. 1 might indicate that Laboratory 21 obtained average results somewhat different from

Table 2. Mean squares from analyses of variance of nicotine by individual laboratories, Study 1

Variation	Degrees of Freedom	Laboratory									
		1	2	4	8	13	15	19	21	24	25
Days	1	0.0526	0.0648	0.0006	0.1823	0.0960	0.2739	0.0397	0.0003	0.0048	0.3610
Ports	19	0.0389	0.0338	0.0487	0.0923	0.0527	0.0470	0.1170	0.0944	0.0432	0.0527
Days × ports	19	0.0297	0.0270	0.0443	0.0907	0.0676	0.0728	0.0738	0.1281	0.0593	0.0743

Table 3. Mean squares from analyses of variance of tar by individual laboratories, Study 1[a]

Variation	Degrees of Freedom	Laboratory 1	2	4	8	13	15	19	21	24
Days	1	5.25	168.51	32.58	155.63	147.07	20.88	232.81	15.25	1.98
Ports	19	15.43	8.31	9.28	9.90	6.87	10.20	16.46	14.74	7.18
Days × ports	19	10.80	8.44	8.62	5.29	11.63	16.59	12.51	21.13	5.53

[a] Laboratory 25 did not report all variables.

Table 4. Mean squares from overall analyses of variance for all 6 measured characteristics on monitor cigarettes, Study 1[a]

Variation	Degrees of Freedom	Cigarette Wt, g	No. of Puffs	Wet Wt TPM, mg	Water, mg	Nicotine, mg	Tar, mg
Laboratories	7	0.077	83.51	263.07	75.84	6.61	122.30
Days in labs	8	0.024	11.27	93.53	11.76	0.08	79.11
Ports	19	0.004	2.36	20.30	2.38	0.08	8.83
Ports × labs	133	0.004	1.62	24.57	4.70	0.06	11.81
Ports × days in labs	152	0.003	0.87	21.97	4.14	0.07	11.11

[a] Laboratories 13 and 25 are omitted. All characteristics are evaluated as averages per 5 cigarettes.

Table 5. Average mg nicotine per 5 cigarettes per lot by laboratories, Study 2

Lab.	Lots 1	2	3	4	5	Av.
1	1.08	3.12	5.17	4.33	6.42	4.02
2	1.14	3.36	5.98	4.83	8.11	4.69
4	0.84	3.12	5.79	4.69	7.76	4.44
8	0.97	3.07	5.20	4.37	7.42	4.21
13	1.04	3.15	5.70	4.74	8.02	4.53
15	1.49	3.52	6.09	5.06	8.06	4.84
19	0.93	3.12	5.77	4.74	7.68	4.45
21	1.62	3.60	6.87	5.31	8.69	5.22
24	1.01	3.24	5.99	4.93	8.27	4.69
25	0.98	2.93	5.45	4.56	7.12	4.21
Average[a]	1.14	3.27	5.86	4.78	7.80	4.57

[a] Omitting Laboratories 13 and 25.

Table 6. Average mg tar per 5 cigarettes per lot by laboratories, Study 2

Lab.	Lots 1	2	3	4	5	Av.
1	22.5	53.1	87.7	71.9	110.5	69.1
2	24.6	52.7	92.1	76.0	128.1	74.7
4	19.9	50.1	89.2	73.0	120.3	70.5
8	21.8	51.1	87.9	72.0	125.0	71.6
13	17.9	44.7	83.1	69.4	123.7	67.8
15	22.5	47.9	92.7	75.5	124.5	72.6
19	20.6	47.7	86.9	70.9	119.6	69.1
21	20.3	48.9	92.3	74.1	127.4	72.6
24	19.8	50.3	93.9	77.5	136.4	75.6
25	Not measured					
Average[a]	21.5	50.2	90.3	73.9	124.0	72.0

[a] Omitting Laboratories 13 and 25.

those of other laboratories, but examination of Fig. 2 and Table 1 indicate that Laboratory 21 got essentially the same results as other laboratories. Without an independent reason (such as existed for Laboratory 13) to conclude that Laboratory 21 was not following the procedure with reasonable care and preciseness it was concluded that this laboratory had, in fact, so followed the procedure and that its results are, therefore, properly a part of the answer about how variable the procedure is.

Examination of Fig. 1, the mean squares in Tables 4 and 9, and especially the components in Tables 10 and 11 indicate that variability in results can be attributed largely to 3 generalized sources: (1) the basic underlying determination error of the procedure, or within-laboratory variability; (2) differences among laboratories, or between-laboratory variability; (3) variability among sources of sample material, or sample variability.

Note in both Tables 10 and 11 that there is, in general, a sizable component at the level of the smallest sampling unit and again at the laboratory level (including laboratory × lot in Table 11). Table 11 (for Study 2) also shows a sizable

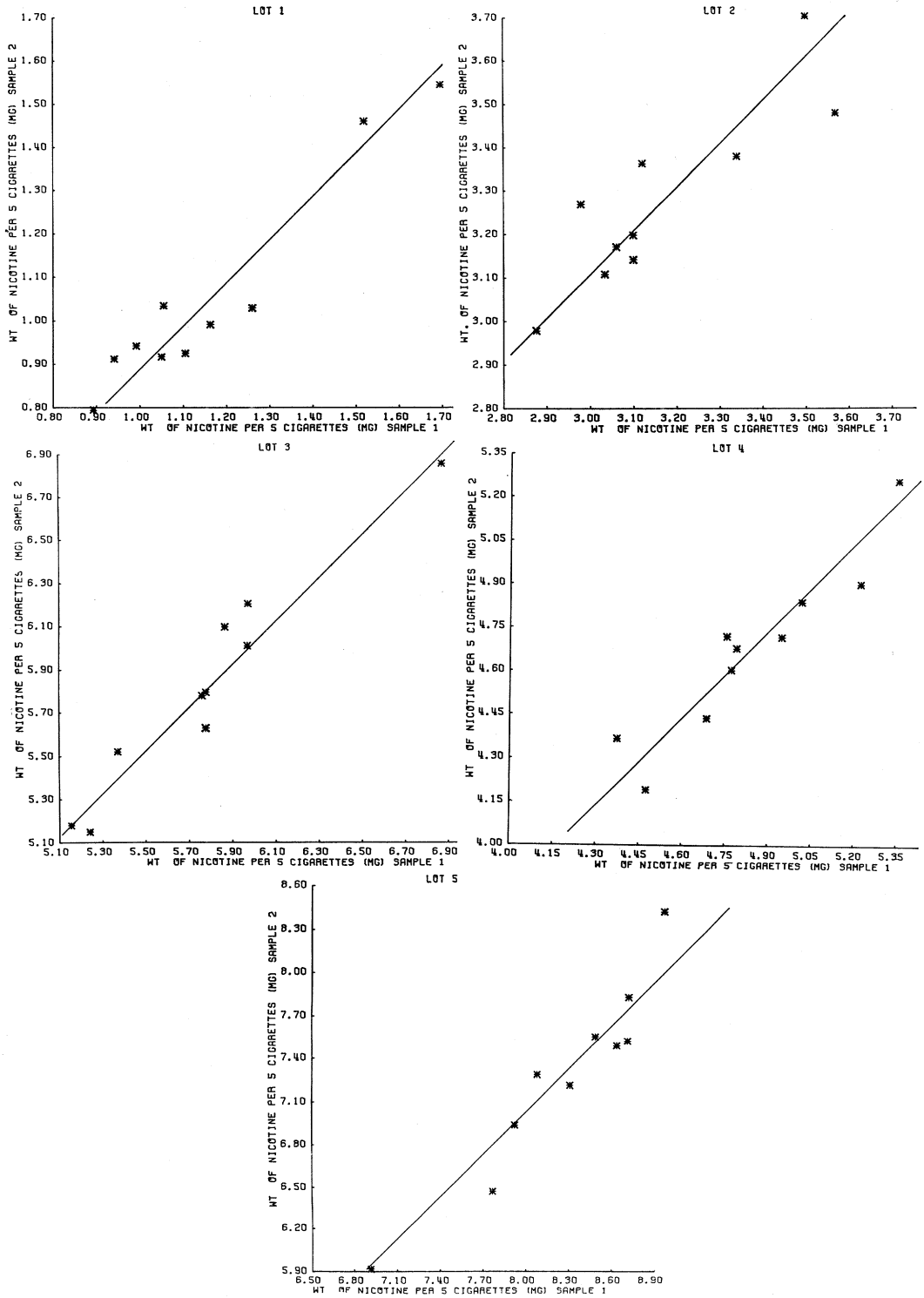

FIG. 1—Two-sample charts: mg nicotine per port, 2 samples each of Lots 1–5.

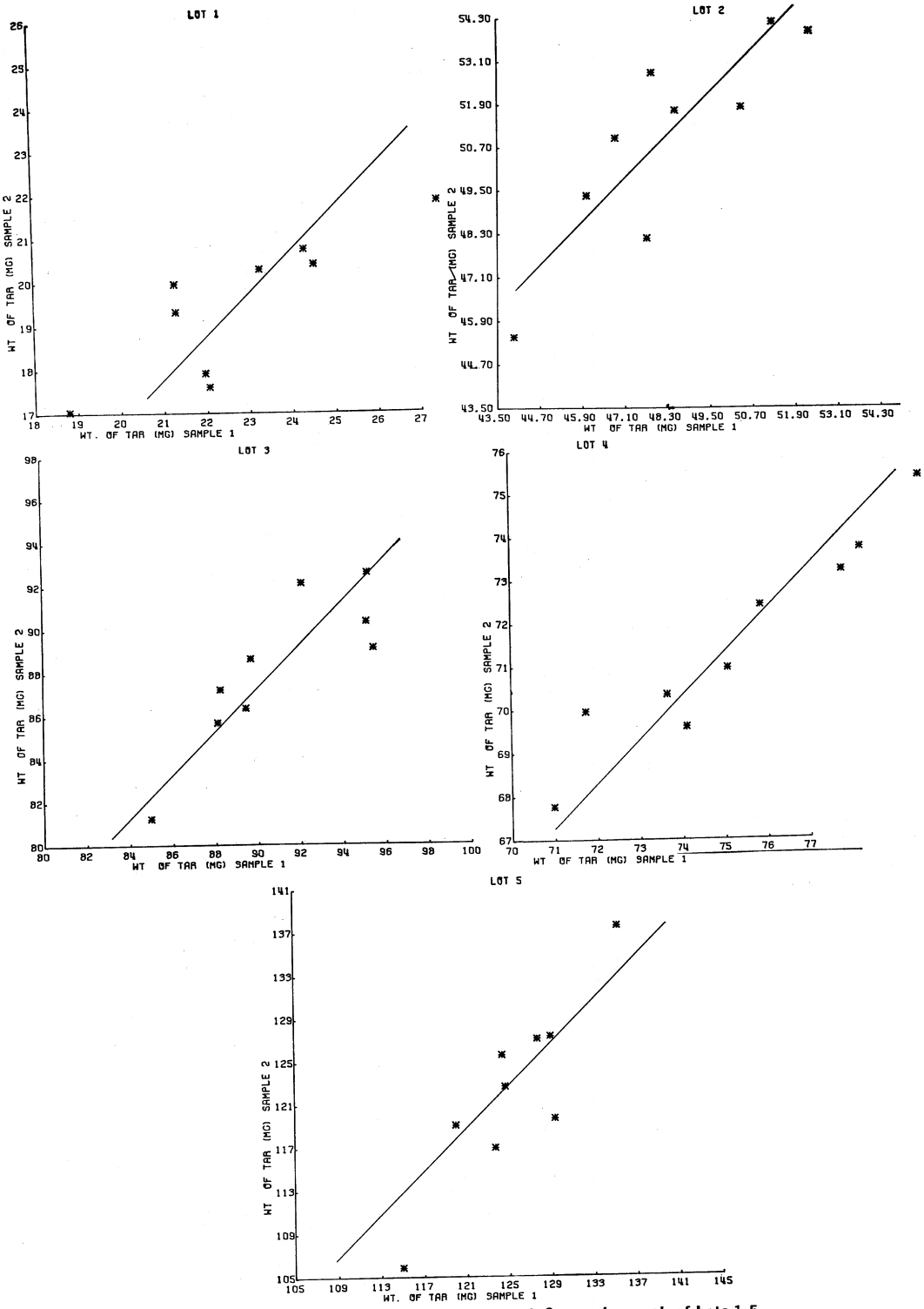

FIG. 2—Two-sample charts: mg tar per port, 2 samples each of Lots 1–5.

Table 7. Mean squares from analyses of variance of tar by individual laboratories, Study 2[a]

Variation	Degrees of Freedom	Laboratories								
		1	2	4	8	13	15	19	21	24
Lots	4	26976.30	36896.00	34867.98	36128.57	38278.48	37336.36	34059.90	40150.65	46717.24
Samples in lots	5	148.95	75.21	109.56	39.13	40.59	181.92	27.32	88.33	88.55
Weeks	1	14.70	1.37	183.52	0.49	10.80	52.40	61.92	52.27	98.83
Days in weeks	4	24.35	34.69	15.69	6.86	53.48	36.80	17.84	24.55	195.54
Lots × weeks	4	5.81	1.08	1.17	1.48	16.24	32.67	12.06	24.05	8.34
Lots × days in weeks	16	71.37	6.89	11.82	6.96	9.41	9.79	15.22	13.21	10.06
Weeks × samples in lots	5	38.38	14.91	8.50	9.58	11.73	3.73	20.62	14.40	18.58
Sample × day in week × lot	20	43.96	9.36	16.50	4.48	19.46	10.79	23.68	23.12	11.45
Ports in sample × day × week × lot	60	24.57	6.36	7.95	6.50	9.85	23.22	20.27	11.66	11.87

[a] Tar reported as mg per 5 cigarettes. Laboratory 25 did not report all variables.

component due to samples, as well as a component associated with laboratory × lot cells (really laboratory × materials). This latter result implies that the procedure is not yet completely stable in the hands of different users, since laboratories may disagree with each other more widely on some lots than on others, an unfortunate circumstance. Fortunately the type of change in disagreement seems to have some regularity and to follow a pattern as shown in Figs. 3 and 4 that another form of analysis might pin it down and suggest a remedy. Sources of variation other than those noted above would not seem to be important.

The general rule for estimating the variance of the mean of a set of observations from the components of variance affecting $\hat{\sigma}^2_{\bar{x}} = [\hat{\sigma}^2$ sampling units/No. sampling units in mean$] + [\hat{\sigma}^2$ subunits/No. subunits in mean$] + [\hat{\sigma}^2$ sub-subunits/No. sub-subunits in mean$] + \ldots$

The above equation may be read in words as follows: The estimated variance of any mean equals the estimated sampling unit component divided by the number of sampling units making up the mean, plus the estimated component due to subsampling units divided by the total number of subsampling units [(number of sampling units) × (number of subunits per unit)] in the mean, plus the estimated component due to sub-subsampling units divided by the total number of sub-subunits [(number of sampling units) × (number of subsampling units per sampling unit) × (number of sub-subsampling units per subsampling unit)] in the mean, etc.

One may use the equation above and the estimated components in Tables 10 and 11 to make estimates about the agreement to be expected among various types or kinds of results (means). More usefully, perhaps, the estimation can be expressed as the magnitude of differences between results that will be equalled or exceeded in only some satisfactorily small percentage of cases. For example, consider 2 evaluations of some sample made in the same laboratory but at different times. How closely should they agree? Turned around, what is the magnitude of the difference that should be equalled or exceeded only 5% of the time in this situation? Such a value might be used by a laboratory as a test of its control over the procedure. If the 2 results should be averaged and compared to the average of 2 other similar results, what is the magnitude

Table 8. Mean squares from analyses of variance of nicotine by individual laboratories, Study 2[a]

Variation	Degrees of Freedom	Laboratory									
		1	2	4	8	13	15	19	21	24	25
Lots	4	99.76	166.36	165.37	138.90	165.72	149.46	158.53	182.09	180.96	132.95
Samples in lots	5	1.33	1.88	1.49	1.18	1.08	1.77	0.86	0.44	1.09	2.15
Weeks	1	0.07	0.04	0.00	0.17	0.25	0.51	0.27	4.10	0.23	0.08
Days in weeks	4	0.19	0.09	0.05	0.24	0.17	0.13	0.24	3.88	0.06	0.06
Lots × weeks	4	0.09	0.15	0.09	0.02	0.09	0.09	0.31	0.65	0.02	0.07
Lots × days in weeks	16	0.20	0.07	0.08	0.09	0.09	0.14	0.17	0.82	0.04	0.02
Weeks × samples in lots	5	0.23	0.09	0.03	0.04	0.01	0.07	0.07	0.28	0.00	0.01
Sample × day in week × lot	20	0.07	0.06	0.06	0.05	0.12	0.07	0.08	0.19	0.07	0.04
Ports in sample × day × week × lot	60	0.14	0.04	0.11	0.05	0.04	0.13	0.10	0.15	0.05	0.04

[a] Nicotine reported as mg per 5 cigarettes.

of the difference that should not be exceeded? It should be obvious that neither sample nor laboratory variability has to be considered in answering these particular questions because the comparisons are always with material obtained from the same sample and with results obtained by the same laboratory, hence there has to be no allowances for uncertainties about samples, nor for laboratories. Consideration must be given, however, to time effects such as days and weeks. The aggregate of these, when combined with the residual component (bottom line) of the analyses of variance, constitute the within-laboratory variance.

Both Study 1 and Study 2 provide estimates of the components needed for estimating the variances of means in situations not involving sample variability or possible laboratory × material interaction (failure of laboratories to maintain constant differences when they evaluate different materials). Only Study 2 provides information about these latter 2 sources of variability.

Tables 12 and 13 show for tar and nicotine, for a number of specified kinds of means, comparisons, and circumstances, the size of difference per port which is expected not to be exceeded more than 5% of the time. All of these estimates hinge on the assumption that we know the true values of the components to substitute into the estimating equations. In practice we do not have such information and we know full well that the values of the components quoted in Tables 10 and 11 are only estimates. They are, however, the very best evidence we have, so we use them. The com-

Table 9. Mean squares from overall analyses of variance for all 6 measured characteristics, Study 2[a]

Variation	Degrees of Freedom	Cigarette Wt, g	No. of Puffs	Wet Wt TPM, mg	Water, mg	Nicotine, mg	Tar, mg
Lots	4	25.331	2573.90	412475.90	5677.08	1229.82	291106.00
Samples in lots	5	0.127	282.64	728.53	77.77	9.27	496.45
Laboratories	7	0.060	130.31	1431.36	151.55	16.93	686.59
Weeks in labs	8	0.019	11.97	59.84	21.15	0.67	58.19
Days in weeks in labs	32	0.022	6.84	57.07	14.30	0.61	44.54
Lots × labs	28	0.039	9.44	437.48	32.17	1.66	289.57
Labs × samples in lots	35	0.020	3.03	64.76	8.86	0.11	37.51
Lots × weeks in labs	32	0.010	1.70	18.90	3.68	0.18	10.83
Lots × days × weeks in labs	128	0.009	2.37	30.70	3.64	0.20	18.16
Samples × weeks in lot × lab	40	0.001	1.83	28.10	2.79	0.10	16.09
Sample × days in weeks in lot × lab	160	0.008	2.24	26.85	2.78	0.08	17.92
Ports in sample × days in weeks in lot × lab	480	0.008	1.86	22.43	2.42	0.10	14.05

[a] Laboratories 13 and 25 omitted. All characteristics evaluated as averages per 5 cigarettes.

Table 10. Components of variance affecting all 6 measured characteristics of monitor cigarettes, Study 1[a]

Variation	Degrees of Freedom	Coefficient	Cigarette Wt		No. of Puffs		TPM Wet Wt		Water Wt		Nicotine Wt		Tar Wt	
			$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%
Laboratories	7	40	0.00125	21.47	1.787	49.69	4.17	13.45	1.588	24.85	0.16344	71.07	1.06	6.67
Days in laboratories	8	160	0.00106	18.25	0.520	14.46	3.58	11.53	0.381	5.96	0.00058	0.25	3.40	21.35
Ports	19	16	0	0	0.046	1.29	0	0	0	0	0.00022	0.10	0	0
Ports × labs	133	2	0.00070	11.95	0.375	10.42	1.30	4.19	0.278	4.36	0	0	0.35	2.19
Ports × days in labs	152	1	0.00282	48.33	0.868	24.14	21.97	70.83	4.143	64.83	0.06572	28.58	11.11	69.79

[a] Laboratories 13 and 25 omitted. All characteristics evaluated as averages per 5 cigarettes.

FIG. 3—Average mg nicotine per lot obtained by individual laboratories versus overall average obtained by all laboratories (omitting Laboratories 13 and 25).

ponents of within-laboratory variability have been evaluated over a much larger number of cases than the components due to samples and laboratories (including laboratory × material or laboratory × lot interactions) and are, therefore, considerably more reliable than the other components. In consequence, estimates about differences that should not be exceeded at some probability level in future samplings should be regarded with considerably more confidence for cases involving only within-laboratory variance than for cases involving either laboratory variability and/or sample variability. Confidence should also increase if estimates based on Study 1 (Table 10) and Study 2 (Table 11) are essentially in agreement.

Tables 12 and 13, in addition to reporting the differences that should not be exceeded for both tar and nicotine, show for tar the estimated variances of the means ($\hat{\sigma}^2_{\bar{x}}$) for the several kinds of sampling and evaluating procedures reported, the components from which these were computed, and the amounts of the several components involved in the variances after division by the appropriate divisor.

The differences that should not be exceeded over 5% of the time are calculated from the

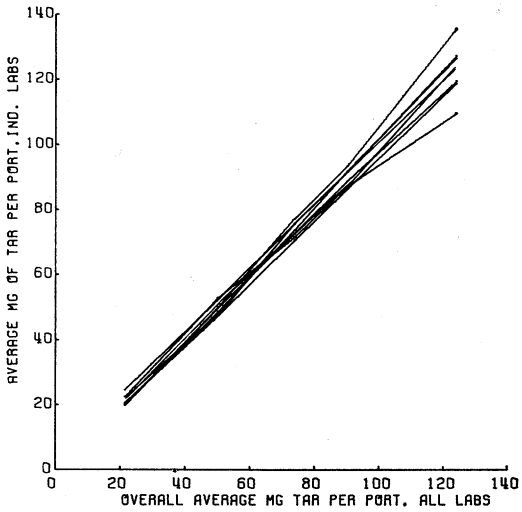

FIG. 4—Average mg tar per lot obtained by individual laboratories versus overall average obtained by all laboratories (omitting Laboratories 13 and 25).

variances of the means involved ($\hat{\sigma}^2_{\bar{x}}$) by calculating the variance of the difference equal $\hat{\sigma}^2_d = 2\hat{\sigma}^2_{\bar{x}}$, taking the square root of this to obtain the standard error of the difference equal $\hat{\sigma}_d = \sqrt{2\hat{\sigma}^2_{\bar{x}}}$ and multiplying this by 2.0, the approximate value of Student's t for the 5% probability level (2-sided) for fairly large numbers of observations. Thus the differences that should not be exceeded more than 5% of the time when comparing 2 means taken under the same sampling and evaluating procedure are calculated as follows:

$$\text{difference} = 2.0\sqrt{2\hat{\sigma}^2_{\bar{x}}}$$

Table 12 reports estimates for nicotine and tar, based on Study 1 (Table 10) and Table 13 reports estimates for nicotine and tar based on Study 2 (Table 11). Ostensibly the situation described in Table 12 as "1 lab, 1 day/lab, 1 port/day, all results from the same sample and laboratory" is the same as that situation described in Table 13 as "1 sample, 1 lab, 1 day/lab, 1 port/day, all means from the same sample and laboratory." It is in each case an estimate of the within-laboratory variation affecting individual results obtained by the procedure. However, the situations are slightly different in that days varied over a longer period of time in Study 2 than in Study 1 with attendant opportunity for greater variations in day effects (including weeks as merely an expression of longer spaced days) both in absolute level and as sample × day and sample × week effects. It is not surprising therefore that the estimates of within-laboratory variation are slightly larger for Study 2 than for Study 1 as the following: standard deviations = 3.85 and 4.20 per port for tar and 0.258 and 0.388 per port for nicotine for Study 1 and Study 2, respectively. These standard deviations per port can be converted to a per cigarette basis by dividing each estimate by 5 to obtain 0.77 and 0.84 for tar and 0.052 and 0.078 for nicotine for Studies 1 and 2, respectively. These compare favorably with within-laboratory values of 1.08 and 0.80 for TPM and 0.061 and 0.054 for nicotine reported by Ogg (1) and with a range of values, 1.2 to 2.8, reported for TPM by Bates *et al.* (4).

For comparing results from different laboratories, the data of this experiment would estimate the standard errors for comparing averages of 8 ports (all in one day) from *different* laboratories to be 0.95 and 0.102 mg per cigarette for tar and nicotine, respectively. These compare with standard deviations of 3.19 and 2.46 for tar and 0.143 and 0.103 for nicotine reported by Ogg (1) for 12 laboratories and with standard deviations of 1.04 and 1.00 for tar and 0.081 and 0.055 for nicotine reported by Ogg (1) for 8 laboratories after deleting 4 obviously different laboratories.

Another comparison of interest might be the average of 2 evaluations (10 cigarettes or 2 ports, days not specified) on a single sample of one lot (brand or source) versus a similar average made by the same laboratory on another lot. The comparison is within laboratories but certainly involves different samples. One might, therefore, conclude that this is the case of "1 sample, 1 lab, 2 days/lab, 1 port per day, different samples, same laboratory" (Table 13). However, the comparison under discussion also involves laboratory × material interaction (labs × lots) which must be included in the estimate of $\hat{\sigma}^2_{\bar{x}}$. Referring to Table 13 and adding 10.492 to the estimate of 15.216 yields $\hat{\sigma}^2_{\bar{x}} = 25.708$. From this the standard error for comparing averages of "duplicate" determinations of tar across lots within any laboratory is $\hat{\sigma}_{\bar{x}} = \sqrt{25.708} = 5.07$ mg/port. Similar calculations for nicotine yield $\hat{\sigma}_{\bar{x}} = \sqrt{0.22900} = 0.478$. In using these estimates of variance and standard deviations it must be remembered that they are constructed from sample estimates of the components and may, therefore, be considerably

Table 11. Components of variance affecting all 6 measured characteristics of the 5 lots of cigarettes, Study 2

Variation	Degrees of Freedom	Coefficient	Cigarette Wt		No. of Puffs		TPM Wet Wt		Water Wt		Nicotine Wt		Tar Wt	
			$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%	$\hat{\sigma}^2$	%
Samples in lots	5	96	0.00114	—	2.9115	—	6.9143	—	0.7183	—	0.09560	—	4.781	—
Laboratories	7	120	0.00031	2.73	0.9796	21.88	11.1365	19.58	1.0360	14.97	0.13543	39.38	5.073	14.58
Weeks in labs	8	60	0	0	0.08553	19.11	0	0	0.1143	1.65	0.00096	0.28	0.227	0.65
Days in weeks in labs	32	20	0.00067	5.90	0.2296	5.13	1.5110	2.66	0.5757	8.32	0.02568	7.47	1.331	3.82
Labs × lots	28	24	0.00079	6.95	0.2615	5.84	15.3697	27.02	1.8680	27.00	0.05786	16.83	10.492	30.15
Labs × samples in lots	35	12	0.00080	7.04	0.0659	1.47	3.0551	5.37	0.5063	7.32	0.00046	0.13	1.632	4.69
Lots × weeks in labs	32	12	0.00010	0.88	0	0	0	0	0.0027	0.04	0	0	0	0
Lots × days in weeks in labs	128	4	0.00020	1.76	0.0325	0.73	0.9630	1.69	0.2154	3.11	0.02656	7.22	0.062	0.18
Samples × weeks in lots × labs	40	6	0	0	0	0	0.2082	0.37	0.0014	0.02	0.00103	0.30	0	0
Samples × days in weeks in lots by labs	160	2	0	0	0.1918	4.28	2.2078	3.88	0.1783	2.58	0	0	1.993	5.55
Ports in samples × days in weeks in lots by labs	480	1	0.00849	74.74	1.8606	41.56	22.4340	39.43	2.4213	34.99	0.09592	27.89	14.051	40.28

Table 12. Estimation from Study 1, for various schemes of sampling and evaluation, of differences in tar and nicotine content that should not be exceeded more than 5% of the time when comparing 2 similarly derived means (with details for tar)

Variation	Estimate of Component, $\hat{\sigma}_i^2$ (Tar)	Amount of Component Appearing in the Mean[a]							
		Same Laboratory(ies)				Different Laboratories			
		1 lab, 1 day/lab, 1 port/day	1 lab, 1 day/lab, 2 ports/day	1 lab, 2 days/lab, 1 port/day	2 labs, 1 day/lab, 1 port/day	1 lab, 1 day/lab, 1 port/lab	1 lab, 1 day/lab, 2 ports/lab	1 lab, 2 days/lab, 1 port/day	2 labs, 1 day/lab, 1 port/day
Laboratories	1.06	0	0	0	0	1.06	1.06	1.06	0.53
Days in laboratories	3.40	3.40	3.40	1.70	1.70	3.40	3.40	1.70	1.70
Ports	0	0	0	0	0	0	0	0	0
Ports × labs	0.35	0.35	0.175	0.175	0.175	0.35	0.175	0.175	0.175
Ports × days in labs	11.11	11.11	5.555	5.555	5.555	11.11	5.555	5.555	5.555
Variance of mean = $\hat{\sigma}_x^2$		14.86	9.13	7.43	7.43	15.92	10.19	8.49	7.96
Difference not exceeded more than 5% of time:									
Tar, mg/port		10.9	8.5	7.7	7.7	11.3	9.1	8.2	8.0
Nicotine, mg/port		0.73	0.52	0.52	0.52	1.36	1.26	1.25	0.96

[a] The amount of any particular component in the mean equals the component divided by the number of such units in the mean. All results from the same sample from the same source.

Table 13. Estimation from Study 2, for various schemes of sampling and evaluation, of differences in tar and nicotine content that should not be exceeded more than 5% of the time when comparing 2 similarly derived means (with details for tar)

		Amount of Component Appearing in the Mean[a]							
		Means from Same Laboratory(ies)				Means from Different Laboratories			
		Means from Same Samples		Means from Different Samples		Means from Same Samples		Means from Different Samples	
	Estimate of Component, $\hat{\sigma}_i^2$ (Tar)	1 sample, 1 lab, 1 day/lab, 1 port/day	1 sample, 1 lab, 2 days/lab, 1 port/day	1 sample, 1 lab, 1 day/lab, 1 port/day	1 sample, 1 lab, 2 days/lab, 1 port/day	1 sample, 1 lab, 1 day/lab, 1 port/day	1 sample, 1 lab, 2 days/lab, 1 port/day	1 sample, 1 lab, 1 day/lab, 1 port/day	1 sample, 1 lab, 2 days/lab, 1 port/day
Samples in lots	4.781	0	0	4.781	4.781	0	0	4.781	4.781
Laboratories	5.073	0	0	0	0	5.073	5.073	5.073	5.073
Weeks in labs	0.227	0.227	0.114	0.227	0.114	0.227	0.114	0.227	0.114
Days in weeks in labs	1.331	1.331	0.666	1.331	0.666	1.331	0.666	1.331	0.666
Labs × lots	10.492	0	0	0	0	10.492	10.492	10.492	10.492
Labs × samples in lots	1.632	0	0	1.632	1.632	1.632	1.632	1.632	1.632
Lots × weeks in labs	0	0	0	0	0	0	0	0	0
Lots × days in weeks in labs	0.062	0.062	0.031	0.062	0.031	0.062	0.031	0.062	0.031
Samples × weeks in lots × labs	0	0	0	0	0	0	0	0	0
Samples × days in weeks in lots × labs	1.933	1.933	0.966	1.933	0.966	1.933	0.966	1.933	0.966
Ports in samples × days in weeks in lots × labs	14.051	14.051	7.026	14.051	7.026	14.051	7.026	14.051	7.026
Variance of mean $= \hat{\sigma}_{\bar{x}}^2$		17.604	8.803	24.017	15.216	34.801	26.000	39.582	30.781
Difference not exceeded more than 5% of time:									
Tar, mg/port		11.9	8.4	13.9	11.0	16.7	14.4	17.8	15.7
Nicotine, mg/port		1.10	0.78	1.40	1.17	1.66	1.47	1.88	1.71

[a] The amount of any particular component in the mean equals the component divided by the number of such units in the mean. All schemes allow that the particular days in a comparison may be more than a week apart

in error. This is especially true for components estimated with few degrees of freedom as was the case for the components due to samples and laboratories. Admitting this deficiency one may proceed to estimate the number of *independent* samples per lot that must be evaluated by this scheme to have a 90% chance of identifying a genuine 5 mg difference per port (1.0 mg difference per cigarette) between lots as significant at the 95% probability level, *if it in fact exists.* Ignoring a very small probability of finding the difference to be significant but in the wrong direction, we have:

$$n = 2(t_\alpha + t_\beta)^2 \hat{\sigma}_{\bar{x}}^2/\delta^2$$

where $t_\alpha = 2.00$, the value of Student's t at the 2-sided 5% level for moderately large degrees of freedom, the significance testing level; $t_\beta = 1.30$, the value of Student's t at the one-sided 10% level for moderately large degrees of freedom, the probability that the difference will be found to be significant if it is in fact present and exactly 5 mg per port; $\hat{\sigma}_{\bar{x}}^2 = (5.07)^2$ for tar and $(0.478)^2$ for nicotine as just estimated; $\delta^2 = (5.0)^2$, the hypothetical difference of 5 mg per port (1 mg per cigarette) on which the problem is based.

The calculation for tar is:

$n = (2)\ (2.00 + 1.30)^2\ (5.07)^2/(5.0)^2 = 23$ independent samples of 10 cigarettes each to be evaluated for each source of interest.

Similarly, for nicotine for $\delta = 0.5$ mg per port (0.1 mg per cigarette), $n = 20$ independent samples of 10 cigarettes, each to be evaluated for each source of interest.

Because of the uncertainties in the estimates of components used in the foregoing estimates of 23 samples for tar and 20 samples for nicotine and because a single sampling scheme must serve for both, it would probably be desirable to state the number of such samples to be from 20 to 25. Remember also that the component for sampling estimated from these studies is for sampling under very favorable conditions and may not extend to sampling commercial material.

Comparisons between the estimates from Studies 1 and 2 of the variability of observations made in different laboratories, as was done for within-laboratory variability, are impossible because Study 1 made no provision for evaluating a possible laboratory × material component, which is shown in Tables 9, 10, and 11 to be significant and quite large for both tar and nicotine. The laboratory × material interation is also shown graphically in Figs. 3 and 4, where it is evidenced by lack of parallelism among lines. The presence of this large component leading to different differences between laboratories according to the material evaluated is judged to be sufficient reason to perform more work on the procedure during the next year and to withhold for the present any recommendation for adoption.

References

(1) Ogg, C. L., *JAOAC* **47**, 356–362 (1964).
(2) Pillsbury, H. C., Bright, C. C., O'Connor, K. J., and Irish, F. W., *JAOAC* **52**, 458–462 (1969).
(3) Youden, W. J., *Statistical Techniques for Collaborative Tests*, AOAC, Washington, D.C., 1967.
(4) Bates, W. W., Griffith, R. B., Harlow, E. S., Senkus, M., and Wakeham, H., *Virginia J. Sci.* **18**, 130–135 (1967).

The recommendation of the Associate Referee was approved by the General Referee and by Subcommittee A and was accepted by the Association. See *JAOAC* **53**, 379 (1970).

This report of the Associate Referee, C. L. Ogg, was presented at the 83rd Annual Meeting of the AOAC, Oct. 13–16, 1969, at Washington, D.C.